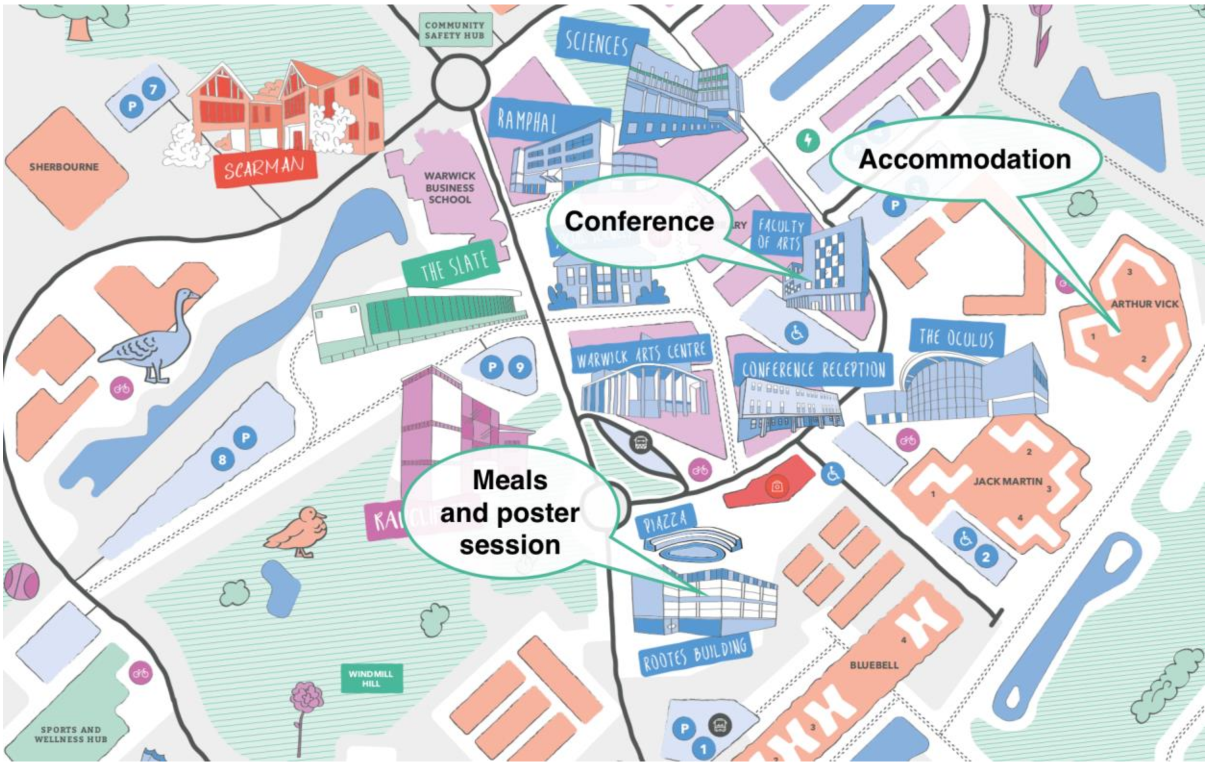
CMS 2024

Computational Molecular Science
Meeting 2024



Accommodation

Conference

Meals and poster session

Wednesday

13:20	**Welcome**	
13:30	**Multiscale Simulations of Biomolecular Motors** (***Sarah Harris***)	
14:20	*Break*	
14:30	DE-FF and MACE-OFF: Data-driven interatomic potentials for molecular simulations (***Daniel Cole***)	Modelling Excited State Processes in Molecular Crystals (***Rachel Crespo-Otero***)
15:00	*Digital Discovery of Semiconducting Polymers: from Chemical Drawing to Electronic Properties* *(**Hesam Makki**)*	*Designer lighting: a radical approach to improving organic light-emitting diodes* *(**Tim Hele**)*
15:20	*Electrostatic embedding of machine learned potentials for accurate and efficient simulation of enzyme catalysis* *(**Elliot Chan**)*	*Ab initio molecular dynamics modelling of the dissociation of hydrofluorocarbon molecules after electron impact in plasma* *(**Dmitry Makhov**)*
15:40	*Coffee Break*	
16:10	Can we learn 'the best' Markov Model? (***Antonia Mey***)	Learning from fundamental surface science for atomic layer deposition – an ab initio endeavour (***Ralf Tonner-Zech***)
16:40	*Combined experimental and computational approaches to characterising the active site structure of cytochrome P460* *(**Kakali Sen**)*	*Atomic-scale insights into frictional energy dissipation mechanisms* *(**Lukas Hörmann**)*
17:00	*The effect of hydration and dynamics on the mass density of single proteins* *(**Matteo Degiacomi**)*	*The impact of Cu amorphous structure on CO2 electrocatalysis : A combined machine learning and DFT modelling approach* *(**Akshayini Muthuperiyanayagam**)*
17:20	*Tunnelling in methylated DNA using NEO-DFT* *(**Juliana Gonçalves de Abrantes**)*	*Computational Insights into the Stability and Phase Transition of Cobalt Oxide Nanoparticles for Fischer-Tropsch Catalysis* *(**Akash Hiregange**)*

Thursday

09:00	**Computational Discovery of New Materials for Singlet Fission in the Solid State** (***Noa Marom***)	
09:50	*Break*	
10:00	*Assessing the accuracy and efficiency of lattice free energy differences obtained from reweighted flow-based probabilistic generative models* *(**Edgar Olehnovics**)*	*Charge transport simulation using kinetic Monte Carlo with fast update rules* *(**Roya Ebrahimi Viand**)*
10:20	*Finding a needle in a needle stack: Leveraging vibrational spectra and machine learning to guide conformational search* *(**Roel van de Ven**)*	*Exploring Triboelectric Charging Mechanisms with Density Functional Theory* *(**James Middleton**)*
10:40	*Coffee Break*	
11:10	Accelerating the prediction of large carbon clusters combining structure search and machine-learning interatomic potentials (***Carla de Tomas***)	Probing Structural Subtleties in Anti-Perovskite Solid Electrolytes (***Karen Johnston***)
11:40	*Thermodynamically Informed Phase Space Exploration for Optimal Autonomous MLIP Dataset Building* *(**Vincent Fletcher**)*	*Determining phase transitions in magnetic materials from first principles: Case study of MnAl* *(**George Marchant**)*
12:00	*Lunch*	
13:30	**More than physics, more than data: integrated machine-learning models for materials** ***(Michele Ceriotti)***	
14:20	*Break*	
14:30	Machine learning of Gaussian basis sets for use in molecular applications (***Grant Hill***)	Exploring the Pt(111)-Electrolyte Interface Under Applied Potentials with Ab Initio Molecular Dynamics (***Clotilde Cucinotta***)
15:00	*MESS: Modern Electronic Structure Simulations* *(**Hatem Helal**)*	*Ab-initio simulation of electron transfer processes of adsorbate-surface systems with ground-state density functional theory* *(**Zsuzsanna Koczor-Benda**)*
15:20	*Machine learning the electric field response of condensed phase systems using perturbed neural network potentials (**Kit Joll**)*	*Computational modelling of oxygen-functionalized graphene: structure-property relationship and application in phosphate sensing* *(**Natalia Martsinovich**)*
15:40	*Coffee Break*	
16:10	**Beyond Crystallinity and Throughput: Machine Learning Accelerated Materials Discovery for Energy Conversion and Storage** *(**Karsten Reuter**)*	
17:00	*Break*	
17:10	**Early Career Networking Event**	
19:00	*Dinner*	

Friday

09:00	**Accelerating the design of Soft Materials using Machine Learning** **(Marjolein Dijkstra)**	
09:50	Break	
10:00	*Atomistic photodynamics simulations for solid-state energy materials* *(**Federico Hernandez**)*	*Machine Learning Order Parameters in Atomistic Systems* *(**Florian Dietrich**)*
10:20	Predicting and Interpreting Femtosecond X-ray spectroscopy (***Tom Penfold***)	Reversible simulation to train classical and machine learning potentials (***Joe Greener***)
10:50	*Coffee Break*	
10:10	*Mapping the landscape of electronic structure theory* *(**Hugh Burton**)*	*Efficient calculation of nucleation rates from equilibrium molecular dynamics and a state correspondence principle* *(**Fabienne Bachtiger**)*
10:30	Surface-Specific Spectroscopy through Machine Learning (***David Wilkins***)	Coatings for Corrosion Protection: Insights from Molecular Simulations (***Flor Siperstein***)
12:00	*Closing Remarks & Lunch*	

Table of Sessions:

Keynote Talks

Multiscale Simulations of Biomolecular Motors

Sarah Harris

University of Sheffield, UK

Molecular motors are vital to living organisms, but their mechanics remains poorly understood. Here I describe a methodology aimed at providing a multiscale computational description of how individual motors perform their function, starting from the atomistic level and moving into the mesoscale using a continuum mechanics representation of proteins.

Computational Discovery of New Materials for Singlet Fission in the Solid State

Noa Marom

Carnegie Mellon University, USA

Intermolecular singlet fission (SF) is the conversion of a photogenerated singlet exciton into two triplet excitons residing on different molecules. SF has the potential to enhance the conversion efficiency of solar cells by harvesting two charge carriers from one high-energy photon, whose surplus energy would otherwise be lost to heat. The development of commercial SF-augmented modules is hindered by the limited selection of molecular crystals that exhibit intermolecular SF in the solid state. Computational exploration may accelerate the discovery of new SF materials. The GW approximation and Bethe–Salpeter equation (GW+BSE) within the framework of many-body perturbation theory is the current state-of-the-art method for calculating the excited-state properties of molecular crystals with periodic boundary conditions. In this talk, I will discuss the usage of GW+BSE to assess candidate SF materials, as well as its combination with low-cost physical or machine learned models in materials

More than physics, more than data: integrated machine-learning models for materials

Michele Ceriotti

École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Machine learning is often used to make ""end-to-end"" predictions, estimating of a property of interest using only a coarse description of the corresponding inputs, in a black-box manner. In contrast, atomic-scale modeling is most useful when it allows gathering mechanistic insights into the microscopic processes that underlie the behavior of molecules and materials.

I will provide an overview of the progress that has been made combining these two philosophies, using data-driven techniques to build surrogate models of the quantum mechanical behavior of atoms, enabling "bottom-up" simulations that reveal the behavior of matter in realistic conditions with uncompromising accuracy.

I will discuss several applications of these ideas, from the calculation of electronic excitations to the design of solid-state electrolyte materials for batteries and high-entropy alloys for catalysis, emphasizing both the accuracy and the interpretability that can be achieved with a hybrid modeling approach.

Beyond Crystallinity and Throughput: Machine Learning Accelerated Materials Discovery for Energy Conversion and Storage

Karsten Reuter

Fritz Haber Institute of the Max Planck Society, Germany

More performant and durable materials are urgently needed to further drive the transition to a sustainable energy system. Unfortunately, accelerated materials discovery is in this field presently still more claim than practical reality. Computational screening approaches hinge on efficient descriptors that only reflect nominal materials properties of the crystalline bulk, simple bulk-truncated surfaces or idealized lattice-matching interfaces. They can thus not account for the substantial, complex and continuous structural, compositional and morphological transitions at the working surfaces or interfaces of catalysts, electrolyzers or batteries. Accelerated experimental discovery in turn still suffers from severe throughput limitations, as easily automatable human steps are rarely limiting the overall workflows. In my talk I will illustrate how modern machine learning (ML) approaches help to overcome these challenges. ML surrogate models, in particular in conjunction with agile active.

Accelerating the design of Soft Materials using Machine Learning

Marjolein Dijkstra

Utrecht University, the Netherlands

Predicting the emergent properties of a material from a microscopic description is a scientific challenge. Machine learning and reverse-engineering have opened new paradigms in the understanding and design of materials. However, this approach for the design of soft materials is highly non-trivial. The main difficulty stems from the importance of entropy, and the ubiquity of multi-scale and many-body interactions. In this talk, I will address questions like: Can we use machine learning to autonomously identify local structures, detect phase transitions, classify phases and find the corresponding order parameters, can we identify the kinetic pathways for phase transformations and can we use machine learning to coarse-grain our models? Finally, I will show in this talk how one can use machine learning to reverse-engineer the particle interactions to stabilize quasicrystals, liquid crystals, and crystals.

Session 1A (Wednesday, 14:30-15:40)

DE-FF and MACE-OFF: Data-driven interatomic potentials for molecular simulations

Daniel Cole

Newcastle University, UK

Drawing on computational methods that are based around training to extensive condensed phase physical property and quantum mechanical datasets, I will describe some of our efforts to design accurate and transferable inter- and intra-molecular potentials, with a view to applications in condensed phase atomistic modelling and computer-aided drug design.

I will explain how recent collaborations with the Open Force Field Initiative enable the development of a fast, accurate alternative to the Lennard-Jones non-bonded potential [1]. With OpenFF, we developed Smirnoff-plugins as a flexible framework to extend the software stack to include custom force field functional forms. We deployed the infrastructure that OpenFF has provided for optimising parameters against condensed phase data, to train a transferable, small molecule force field based on a double exponential functional form (DE-FF). The automated framework allowed us to train and test a full small molecule force field in just a matter of weeks (as opposed to many years for traditional force fields), with promising accuracy in the condensed phase.

Finally, I will describe MACE-OFF23, a transferable force field for organic molecules created using state-of-the-art machine learning technology and first principles reference data. MACE-OFF23 demonstrates the remarkable capabilities of short-range models by accurately predicting a wide variety of gas and condensed phase properties of molecular systems, including dihedral scans, descriptions of molecular crystals and liquids, and even properties of a solvated small protein.

[1] Horton, J., et al. Digital Discovery, 2023, 2, 1178. https://doi.org/10.1039/D3DD00070B

[2] Kovács, D., Moore, J.H., et al. (2023). arxiv, https://arxiv.org/abs/2312.15211"

Digital Discovery of Semiconducting Polymers: from Chemical Drawing to Electronic Properties

Hesam Makki

University of Liverpool, UK

Coauthors: Hesam Makki, Colm Burke, Alessandro Troisi

Department of Chemistry and Materials Innovation Factory, University of Liverpool, Liverpool L69 7ZD, U.K.

Semiconducting polymers (SCPs) are the core material for flexible electronics with many applications including organic photovoltaics, bioelectronics, and FETs. The modular approach to their synthesis lends itself to a natural approach to molecular design, namely the selection of a sequence of conjugated monomers and sidechains in the repeat unit structure for a targeted application. So far, materials development in the field has progressed largely by trial and error and a generic approach that provides the necessary design principles for the desired properties in SCPs is in extremely high demand. Existing computational methods are used to model SCPs, but there are several challenges that have inhibited the expected progress in the field thus far. As such, compared to the volume of new materials introduced every year there is a very limited number of SCPs with reliable atomistic models since accurate simulations are reserved for just a few polymers per investigation. Expediting polymer modelling is achieved by changing the current approach in the modelling of polymers: instead of building a model for a material, one needs to build a protocol for a material class.

In this research, we introduce our in-house software which enables digital discovery of SCPs [1]. By employing this software, we generate a library of (tens of) existing and hypothetical SCPs by expediting the modelling processes through developing a standard automated workflow. Our computational approach i) employs a multiscale resolution method (from sub-atomic to mesoscale), ii) enables speedy model generation, iii) develops a high-throughput QM/MM method to establish (currently lacking) structure-property relationship dataset for this material class, and iv) generates new design rules for future generations of SCPs. This software not only enables unravelling the charge transport mechanisms of benchmark polymers [2], but also facilitates formulation of many hypothetical SCP structures with promising electronic properties as future alternatives. In this paper, we summarise the results of modelling of more than 100 (existing and hypothetical) SCPs and demonstrate several examples of design principles obtained by such study.

References:

1. Burke C, Makki H, Troisi A. J. Chem. Theory Comput. 2024, 20, 9, 4019–4028.

2. H. Makki, C. Burke, A. Troisi, J. Phys. Chem. Lett. 2023, 14, 39, 8867–8873 .

Electrostatic embedding of machine learned potentials for accurate and efficient simulation of enzyme catalysis

Elliot Chan

University of Bristol, UK

Coauthors: Elliot Chan1, Kirill Zinovjev2. Lester Hedges3, Christopher Woods3, Marc van der Kamp1

1: School of Biochemistry, University of Bristol, Biomedical Sciences Building, University Walk, Bristol BS8 1TD, UK

2: Departamento de Química Física, Universidad de Valencia, 46100, Burjassot

3: Research Software Engineering, Advanced Computing Research Centre, 31 Great George Street, Bristol BS1 5QD, UK

To simulate reactions in large biomolecular systems, Quantum Mechanics/Molecular Mechanics (QM/MM) is typically used, treating just the region where the reaction occurs with QM and the rest with MM. Whilst this is the state-of-the-art for simulating enzyme reactions, the speed is primarily determined by the QM method used, drastically limiting accessible timescales and sampling.

An emerging solution is to employ machine learning potentials (MLPs) that predict energies of atomic arrangements at a fraction of time needed for QM. Pure MLPs would still be slower than MM, but "ML/MM" simulations would offer significant speed-up over QM/MM. One cannot simply replace QM with an MLP: electrostatic embedding, crucial for capturing catalytic effects, would not work. Due to the absence of electronic density in MLPs, the influence of MM partial charges cannot be taken into account.

Here, we present an electrostatic machine learning embedding scheme (EMLE), in which generic in vacuo MLPs are implemented in an efficient ML/MM simulation, with EMLE predicting the response of the ML region to the MM region (https://github.com/chemle). We test EMLE on simulations of the Diels-Alder reaction catalysed by the natural Diels-Alderase AbyU. An Atomic Cluster Expansion (ACE) MLP was trained on the reaction in the gas phase, learning energies based on the M06-2X functional. Umbrella sampling of the reaction with ML/MM(EMLE) and this MLP accurately reproduced the free energy profile from QM/MM using M06-2X with significant speed up (up to 285 times). Further, tests on different enzyme-substrate poses reveals that 'mechanical embedding' (where fixed point charges are used for the MLP to provide electrostatic interactions with the MM region) is unable to distinguish between the known reactive pose and alternative non-reactive poses in AbyU. We thus demonstrate that efficient ML/MM reaction modelling is possible, with EMLE crucial to capture enzyme catalysis accurately.

Session 1B (Wednesday, 14:30-15:40)

Modelling Excited State Processes in Molecular Crystals

Rachel Crespo Otero

University College London (UCL), UK

Coauthors: Federico Hernandez (1,2), Michael Ingham(1) and Amir Sidat (3)

1 University College London

2 University of Bristol

3 Queen Mary University of London

The study of photoexcitations in molecular aggregates faceSs the twofold problem of the increased computational cost associated with excited states and the complexity of the interactions among the constituent monomers. A mechanistic investigation of these processes requires the analysis of the intermolecular interactions, the effect of the environment, and 3D arrangements or crystal packing on the excited states. A considerable number of techniques have been tailored to navigate these obstacles; however, they are usually restricted to in-house codes and thus require a disproportionate effort to adopt by researchers approaching the field. In this talk, I will describe some techniques implemented in our group to investigate excited state processes in molecular crystals, considering electrostatic embedding approaches and their use for studying nonadiabatic dynamics. I will showcase a recent application for studying singlet fission in the pentacene crystal.

Designer lighting: a radical approach to improving organic light-emitting diodes

Tim Hele

University College London, UK

Coauthors: Timothy J. H. Hele, James D. Green, Emrys W. Evans, Richard H. Friend, and Feng Li

1, 2. University College London, UK

3. University of Swansea, UK

4. Cavendish Laboratory, University of Cambridge, UK

5. Jilin University, China

In 2018 a team including myself discovered the world's most efficient deep-red light-emitting diode and published this in Nature [1]. Surprisingly this was not based on a closed-shell molecule but on a radical emitter with an unpaired electron. Previously radicals had been considered to be too unstable or non-emissive for serious consideration in most photochemical applications, and this discovery opened a new field to exploration. In 2020 we then proceeded to show, using theoretical arguments, why many previously-synthesised radicals were not emissive and provided design rules for determining, prior to computation or synthesis, whether a radical was likely to emit [2]. We then demonstrated how a triplet molecule can give its energy to a radical which then emits, improving conventional organic light-emitting diodes (OLEDs) based on a closed-shell emitter [3]. In this talk I will give an overview of radical OLEDs and cover recent advances such as the use of inverse molecular design [4] and the fast, accurate and spin-pure calculation of the excited states of radicals [5].

[1] Nature 563 536–540 (2018)

[2] Nature Materials 19, 1224-1229 (2020)

[3] Nature Communications 13 ,2744 (2022)

[4] J. Chem. Phys. 156, 180901 (2022)

[5] J. Chem. Phys. (accepted) (2024) https://arxiv.org/abs/2312.09013

Ab initio molecular dynamics modelling of the dissociation of hydrofluorocarbon molecules after electron impact in plasma

Dmitry Makhov

University of Leeds, UK

Coauthors: Dmitry V. Makhov, Ryan Brook, Dmitrii V. Shalashilin

School of Chemistry, University of Leeds, Leeds, LS2 9JT, United Kingdom

Ab Initio Multiple Cloning approach (AIMC) [1, 2] is used to simulate the process of dissociation for hydrofluorocarbon molecules in low triplet states excited by electron impact in plasma. The interest in the dissociation of hydrofluorocarbons in plasma is motivated by their role in plasma etching in microelectronic technologies. Although AIMC method was originally developed for the simulation of nonadiabatic dynamics of excited molecules in singlet states in photochemistry, it can equally be used for the dynamics of molecules in low-energy triplet states produced by electron impact. Our study concentrates on these triplet states because they are the lowest electronic states often separated from other excited states by a substantial energy gap. Calculations show that for many molecules higher triplet states relax extremely quickly into the lowest triplet state, such that the simplest ab initio MD on the lowest triplet state seems to give a reasonable estimate of the reaction channels branching ratios [3]. The calculations appear to yield very simple rules that can be used to predict dissociation channels.

[1] D.V. Makhov, W.J. Glover, T.J. Martinez, D.V. Shalashilin, J. Chem. Phys. 141 (2014) 054110.

[2] D.V. Makhov, C. Symonds, S. Fernandez-Alberti, D.V. Shalashilin, Chem. Phys. 493 (2017) 200.

[3] D.V. Makhov, G. Armstrong, H. Chuang, H. Ambalampitiya, K. Lemishko, S. Mohr, A. Nelson, J. Tennyson, and D.V. Shalashilin, J. Phys. Chem. Lett., 15 (2024) 3404.

Session 2A (Wednesday, 16:10-17:40)

Can we learn ‘the best’ Markov Model?

Antonia Mey

University of Edinburgh, UK

Coauthors: Antonia Mey1, Yanchen Zhu1, Robert Arbon1

1 EaStCHEM School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ, U.K.

Markov state models (MSM) have been a widely used tool for the analysis of conformational dynamics of proteins, including protein folding. As it is a statistical and machine learning (ML) model certain choices must be made in constructing an MSM. These choices, or hyperparameters, are often chosen by expert judgement or maximizing variational scores such as the VAMP-2 score. Modern ML and statistical pipelines often use automatic hyperparameter selection techniques ranging from the simple: choosing the best score from a random selection of hyperparameters to the complex: optimization via e.g., Bayesian optimization. Can we use these out-of-the-box optimization techniques to select ‘the best’ MSMs automatically?

In this talk, I will walk you through how we systematically built and assessed over a million MSMs for protein folding models and why model observables and variational scores should only be used to guide model selection. Furthermore, I will show some recent highlights of how we can use MSM to extract longer-than-millisecond timescale dynamics from around 1 ms aggregated simulation time of a kinase.

Combined experimental and computational approaches to characterising the active site structure of cytochrome P460

Kakali Sen

UKRI-STFC, UK

Coauthors: Kakali Sen1, Aditya G. Rao2, Hans Pfalzgraf3, Hannah Adams4, You Lu1, Chin Yong1, Richard Strange4, Marcus Edwards4, Adrian Mulholland2, Michael Hough3, Thomas Keal1

1 Scientific Computing Department, STFC Daresbury Laboratory, Warrington, Cheshire WA4 4AD, United Kingdom

2 School of Chemistry, Centre for Computational Chemistry, University of Bristol, Bristol Bs8 1TS, United Kingdom

3 Diamond Light Source Ltd, Harwell Science and Innovation Campus, Fermi Avenue, Didcot, OX11 0DE, United Kingdom

4 School of Life Sciences, University of Essex, Colchester, Essex CO4 3SQ, United Kingdom

The nitrification pathway of the global nitrogen cycle converts ammonia to nitrate via a series of oxidation reactions that produces several intermediates, including NO, N2O, NH2OH and NO2-, that are toxic pollutants and/or contribute to climate change. The first step of the nitrification process oxidises ammonia to NH2OH, which is then itself oxidised to the next intermediate (N2O or NO). One of the two enzymes capable of this reaction is the cytochrome P460s (CytP460). These enzymes have an unusual heme chemically modified through interaction with a Lys amino-acid side chain in a post-translational modification i.e. a protein-heme crosslink is formed. Understanding how bacteria use CytP460 to convert NH2OH to the greenhouse gas N2O is of major agricultural and environmental significance and such cross-links play a role in dictating the activity of this enzyme.

Prior structural characterisation of CytP460 reveals it as a dimeric protein with each monomer containing a single heme P460 with a Lys crosslink.1, 2 In our current combined experimental and computational work on P460s, crystal structures of P460 from M. capsulatus (McP460) indicate that this enzyme can form more than one crosslink between the Lys residue and the heme. Utilising hybrid QM/MM studies we are simulating the UV-Vis spectra to uncover the electronic structure and the nature of these multiple cross-link species in McP460. Furthermore, comparison of the spectra with P460 from N. europaea (NeP460)3, which forms just one crosslink with the Lys residue will provide some understanding how the differences in the active site of these proteins (these two proteins bear 52% sequence similarity but are highly structurally homologous) results in such a complex crosslink, though both perform the same transformation of NH2OH to N2O. Our work also provides insights into the structure of the experimentally challenging ferrous form of cytochrome P460.

1. Adams et al Chem Sci 2019, 10, 3031.
2. Pearson et al Biochemistry 2007, 46, 8340.
3. Smith et al Biochemistry 2018, 57, 334."

The effect of hydration and dynamics on the mass density of single proteins

Matteo Degiacomi

Durham University, UK

Coauthors: Cameron C. W. McAllister, Elizabeth H. C. Bromley, Matteo T. Degiacomi

Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK

The density of a protein molecule, typically consider as a constant, is a key property used in a variety of experimental techniques, from X-ray crystallography to ultracentrifugation. We have developed an accurate molecular dynamics-based method for determining protein mass density that directly incorporates the influence of hydration. We find that different proteins have different density, dependent on a range of physical parameters associated with amino acid composition. Utilising correlations with sequence-derived characteristics, we then develop a random forest regressor which can be used to accurately predict protein density from a given amino acid sequence. Finally, using titin and Bovine Pancreatic Trypsin Inhibitor as prototypical examples, we show that individual proteins can occupy states with close but distinguishable densities, thereby demonstrating that proteins may be Raman active at low frequencies.

Tunnelling in methylated DNA using NEO-DFT

Juliana Gonçalves de Abrantes

University of Surrey, UK

Coauthors: Juliana G. de Abrantes (1), Louie Slocombe (2), Brendan Howlin (1), Marco Sacchi (1)

(1) University of Surrey

(2) Arizona State University

Methylation of DNA nucleobases is a naturally occurring process in living organisms. Usually, it functions as a gene regulation marker [1], and is connected to inheritable epigenetic effects [2].

Methylation of Guanine in the O6 position disrupts the hydrogen bonding between pairing bases [3], which might affect reactions and dynamics in DNA. In this work, we investigate how this base modification affects the Double Proton Transfer (DPT) between methyl-Guanine and Cytosine, originating tautomers that can lead to point mutations in DNA.

Previous results show that DPT can happen through quantum tunnelling [4], and that the tunnelling probability during the replication process is affected by strand separation and the presence of helicase [5]. If the tautomeric species survive the replication machinery, the point mutations will persist in the organism and may lead to the development of diseases like cancer.

We compare results obtained from Density Functional Theory (DFT) calculations to the Nuclear-Electronic Orbital (NEO-DFT) method [6], where the protons involved in the transfer are treated at the same Quantum-Mechanical level as the electrons of the system, which gives us a better theoretical treatment of the tunnelling phenomenon. Indeed, we find that the effective energy barriers calculated by NEO-DFT are around 20% lower than the classical barriers, indicating that the tunnelling probability is higher than previously estimated.

[1] R. Murín, M. Abdalla, N. Murínová, J. Hatok, and D. Dobrota. The Metabolism of 5-methylcytosine Residues in DNA. Physiological Research, 2018.

[2] S. McGraw and S. Kimmins. Inheritance of epigenetic DNA marks studied in new mouse model. Nature, 2023.

[3] T. Spratt. Structure of the hydrogen bonding complex of O6-methylguanine with cytosine and thymine during DNA replication. Nucleic Acids Research, 1997.

[4] L. Slocombe, J. S. Al-Khalili, and M. Sacchi. Quantum and classical effects in DNA point mutations: Watson–Crick tautomerism in AT and GC base pairs. Physical Chemistry Chemical Physics, 2021.

[5] L. Slocombe, M. Winokan, J. S. Al-Khalili, and M. Sacchi. Proton transfer during DNA strand separation as a source of mutagenic guanine-cytosine tautomers. Communications Chemistry, 2022.

[6] Qi Yu and S. Hammes-Schiffer. Nuclear-Electronic Orbital Multistate Density Functional Theory. The Journal of Physical Chemistry Letters, 2020.

Session 2B (Wednesday, 16:10-17:40)

Learning from fundamental surface science for atomic layer deposition – an ab initio endeavour

Ralf Tonner-Zech

Leipzig University, Germany

Wilhelm-Ostwald-Institut für Physikalische und Theoretische Chemie, Leipzig University, Linnéstraße 2, 04103 Leipzig, Germany

Atomic Layer Deposition (ALD) emerges as an important technique in addressing the evolving complexities of microelectronics fabrication, offering the potential of atomically precise material construction. But as every practitioner of the method knows, the reality is much more complex and many physical and chemical effects have to be considered, tested and investigated to arrive at the goal of a targeted materials synthesis. A current forefront challenge is achieving area-selective ALD (AS ALD), aiming for selective material growth on target surfaces, with small molecule inhibitors (SMIs) presenting a promising strategy to prevent deposition on non-growth areas.

The nature of the process requires an in-depth understanding of the underlying surface chemistry as well as a mechanistic understanding of the adsorption and decomposition reactions that can happen in the SMI as well as the ALD process steps. Our latest research shines a light on SMI-based AS ALD, offering insights into experimental results and pushing the boundaries toward predictive computational analysis.1–4

At its core, ALD is governed by surface chemistry principles, allowing extrapolations from fundamental studies on molecular-surface interactions, often performed in more experimentally controlled settings. Extensive research into the reactivity of organic adsorbates on semiconductor surfaces has unearthed key reaction mechanisms, bridging insights reactivity known from molecular chemistry.5,6 For metal surfaces, while a local view on chemical bonding may fall short, the interplay with organic adsorbates is still within reach of computational approaches. Our findings include investigations into the stability of the longest known acene chain on Cu(111)7 and the behavior of non-alternant aromatic compounds on metallic substrates.8

One focus of the work is based around understanding the electronic structure and revealing the underlying chemical driving forces using our energy decomposition method for extended systems (pEDA) which allows for quantitative analysis.[6]

1. Yarbrough, J. et al. Area-Selective Atomic Layer Deposition of Al2O3 with a Methanesulfonic Acid Inhibitor. Chem. Mater. 35, 5963–5974 (2023).

2. Yarbrough, J. et al. Tuning Molecular Inhibitors and Aluminum Precursors for the Area-Selective Atomic Layer Deposition of Al2O3. Chem. Mater. 34, 4646–4659 (2022).

3. Oh, I.-K. et al. Elucidating the Reaction Mechanism of Atomic Layer Deposition of Al2O3 with a Series of Al(CH3)xCl3–x and Al(CyH2y+1)3 Precursors. J. Am. Chem. Soc. 144, 11757–11766 (2022).

4. Zoha, S., Pieck, F., Gu, B., Tonner-Zech, R. & Lee, H.-B.-R. Organosulfide Inhibitor Instigated Passivation of Multiple Substrates for Area-Selective Atomic Layer Deposition of HfO2. Chem. Mater. (2024) doi:10.1021/acs.chemmater.3c02525.

5. Pecher, L., Laref, S., Raupach, M. & Tonner, R. Ethers on Si(001): A Prime Example for the Common Ground between Surface Science and Molecular Organic Chemistry. Angew. Chem. Int. Ed. 56, 15150–15154 (2017).

6. Pecher, L. & Tonner, R. Deriving bonding concepts for molecules, surfaces, and solids with energy decomposition analysis for extended systems. WIREs Comput. Mol. Sci. 9, e1401 (2019).

7. Ruan, Z. et al. Synthesis of Tridecacene by Multistep Single-Molecule Manipulation. J. Am. Chem. Soc. 146, 3700–3709 (2024).

8. Klein, B. P. et al. Molecular Topology and the Surface Chemical Bond: Alternant Versus Nonalternant Aromatic Systems as Functional Structural Elements. Phys. Rev. X 9, 011030 (2019).

Atomic-scale insights into frictional energy dissipation mechanisms

Lukas Hörmann

University of Warwick, UK

Coauthors: Lukas Hörmann1; Shinjae Nam2; Alfred J. Weymouth2; Franz J. Gießibl2; Reinhard J. Maurer1

1 University of Warwick, Coventry, UK; 2 University of Regensburg, Regensburg, Germany

Friction causes significant energy loss in any moving mechanical device. As the miniaturisation of devices reaches the quantum limit, so do dynamical dissipation processes. Fundamentally quantum mechanical mechanisms govern friction at the atomic-scale. We account for all relevant quantum mechanical effects, such as charge transfer, or van der Waals interactions, by employing density functional theory, and machine learning and use the probe particle model to simulate energy dissipation. Using this approach, we investigate dynamic friction at the natural limit of a singular atom moving a single chemical bond. This enables us to explore how the local bonding environment of the underlying sample surface affects the energy dissipation that a probe particle experiences. Focussing on the example of a CO-functionalised lateral force microscope that measures frictional energy dissipation above various organic adlayers on Cu(111), we find strong correlations between the local bonding environment and the energy dissipation. Our findings capture the qualitative trends found in experiment. Finally, we present a mechanistic interpretation of our findings that provides insights into the underlying physics of atomic/lateral force microscopy measurements.

The impact of Cu amorphous structure on CO2 electrocatalysis : A combined machine learning and DFT modelling approach

Akshayini Muthuperiyanayagam

Queen Mary, University of London, UK

Coauthors: Akshayini Muthuperiyanayagam 1, Dr Devis Di Tommaso 1

1 Department of Chemistry, School of Physical and Chemical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, United Kingdom

Amorphous materials hold significant promise for enhancing electrocatalytic CO2 reduction (CO2R) performance, but their intricate structures present challenges in understanding their behaviour. We present a computational investigation combining machine learning force fields and DFT calculations to explore amorphous copper (Cu) as a potential catalyst for CO2R to C1 and C2 products. Our study reveals that amorphous Cu surfaces, compared to crystalline counterparts, offer a wider range of coordination sites, leading to a multitude of active centres for CO2 adsorption. Notably, some investigated amorphous surfaces spontaneously activate CO2, demonstrating their potential for efficient conversion. Furthermore, the intermediates of the CO2R on these surfaces exhibit enhanced stability, translating to lower overpotentials and improved selectivity. The next step in this study will explore how amorphization can be combined with metal doping to achieve synergistic effects in amorphous Cu-based catalysts. This work paves the way for further research and development in using amorphous Cu-based catalysts for sustainable CO2 conversion technologies.

Computational Insights into the Stability and Phase Transition of Cobalt Oxide Nanoparticles for Fischer-Tropsch Catalysis.

Akash Hiregange

Cardiff University, UK

Coauthors: Akash Hiregange*, Pavel Stishenko, Andrew Logsdail, James Paterson, Sarah Haigh

School of chemistry, Cardiff University, Park place, Cardiff, CF10 3AT, UK

The Fischer-Tropsch (FT) reaction, involving the conversion of syngas (CO and H2 mixture) to liquid hydrocarbons, is important in producing sustainable hydrocarbon fuels. The energy barrier for CO hydrogenation in FT reaction necessitates the use of transition metal catalysts to make the process feasible.[1] Cobalt-containing metal oxides, dispersed on TiO2 support, have emerged as a precursor of metallic catalysts for Fischer-Tropsch synthesis, along with Mn promoters to improve the selectivity.2 In the Mn-promoted Co/TiO2 FT catalyst, an increase in Mn2+ loading enhances the alcohols (oxygenates) selectivity.[2] Mn addition results in smaller nanoparticles, enhanced dispersion of cobalt oxide over the support and a phase transition from spinel Co3O4 to rocksalt (RS) CoO, indicating transition metal (Co, Mn and Ti) interactions. However, nature of the interactions that drive the phase transition of cobalt oxide in the catalytic precursor, remains elusive.

Density functional theory (DFT) has been used in this work to investigate the stability of Co3O4 and CoO nanoparticles as a function of particle size to rationalize the cobalt oxide phase transition observed in experiments for smaller particle sizes. Slab calculations of RS CoO reveal that the high miller index facets are more stable than the low miller index facets. For Co3O4, the (001) facet was found to be the most stable. The relative stabilities of the surface facets were used to predict the equilibrium nanoparticle morphology via Wulff constructions. Subsequent analysis reveals that the phase transition between RS CoO and Co3O4 occurs at a critical size of ~3.6 nm. Notably, CoO proves to be more stable for particle sizes less than 3.6 nm. These findings concretize the experimental observation of phase transition of cobalt oxide nanoparticles by offering insights about nanoparticle stability over a range of particle.

References

[1] Z. Qinghong, J. Kang, Y. Wang. ChemCatChem, 2(9), 1030-1058 (2010)

[2] J. Paterson, M. Peacock, R. Purves, R. Partington, K. Sullivan, G. Sunley, J. Wilson. ChemCatChem, 10(22), 5154-5163 (2018)

Assessing the accuracy and efficiency of lattice free energy differences obtained from reweighted flow-based probabilistic generative models

Edgar Olehnovics

University College London, UK

Coauthors: Edgar Olehnovics, Michael Shirts, Matteo Salvalaglio

1 University College London, 210B Torrington Place, London, WC1E 7JE

2 University of Colorado Boulder

Solid forms of most active pharmaceutical ingredients exhibit polymorphism. Polymorphism refers to the ability of a molecule to crystallise into more than one type of thermodynamically unique and stable crystal. The relative stability of different polymorphs depends on physical conditions such as temperature and pressure and is quantified by free energy differences. Free energies at finite temperature and pressure, can be accurately estimated by computational methods such as free energy perturbation (FEP). Classically, FEP methods rely on explicitly sampling configurations from multiple unphysical intermediate states to statistically connect the physical metastable states of interest. Modern, targeted FEP methods (TFEP) instead augment FEP with invertible coordinate transformations (maps), carefully chosen to maximise the necessary overlaps, without sampling any intermediate states. In complex systems, this remapping approach is only possible with machine learning. Probabilistic generative models (PGMs) based on normalising-flow architectures can thus make it easier than ever before to obtain arbitrarily complicated maps needed for TFEP. In this contribution, we discuss the robustness of PGMs when reweighted via BAR or MBAR for calculating conformational free energies of isolated molecules. We then extend the insight obtained in this framework to efficiently compute the lattice-free energies of simple crystal supercells (e.g. ice Ih and Ic). These studies are stepping stones in rendering FEP estimates more easily accessible and computationally efficient.

Finding a needle in a needle stack: Leveraging vibrational spectra and machine learning to guide conformational search

Roel van de Ven

Radboud University, NL

Coauthors: Roel van de Ven, Daria Ruth Galimberti

Radboud university, Houtlaan 4, 6525 XZ Nijmegen

The standard approach for identifying the conformer corresponding to an experimental infrared (IR) spectrum involves exploring the potential energy surface (PES), locating the minima and calculating the spectra of each conformation until a match is found. This sampling approach is highly inefficient because it requires numerous frequency calculations on structures that, in the end, are discarded. It becomes unfeasible for larger and larger molecules since the number of minima on the PES will increase exponentially with the number of atoms.

I will present a new methodology to cleverly direct the conformational search using molecular dynamics (MD) simulations informed by an on-the-fly machine learning (ML) model based on generated vibrational spectroscopy data during the search. Our program starts from a random position on the PES and the machine learning model learns on the fly how to modify the molecular conformation in a way to improve the spectral match until the correct conformation is found. The inputs of the ML model are the spectral similarity vectors, i.e., vectors describing the difference between the current and target spectra, and contain all information about intensity and frequency (mis)matches. The outputs of the ML model are proposed biases that should be applied in the next MD simulation to push the molecule to the desired conformer.

I will critically discuss the results we obtained on a test set of calculated spectra generated for the diphenylglycine molecule.

Accelerating the prediction of large carbon clusters combining structure search and machine-learning interatomic potentials

Carla de Tomas

Imperial College London, UK

Coauthors: B. Karasulu [1], J.M. Leyssale[2], P. Rowe[3], C. Weber[4], C. de Tomas[5]

[1] Warwick University, [2] University of Bordeaux, [3] University College London, [4] King's College London, [5] Imperial College London

From as small as single carbon dimers up to giant fullerenes or amorphous nanometer-sized particles, the large family of carbon nanoclusters holds a complex structural variability that increases with cluster size. Capturing this variability and predicting stable allotropes remains a challenging modelling task, crucial to advance technological applications of these materials. While small cluster sizes are traditionally investigated with first-principles methods, a comprehensive study spanning larger sizes calls for a computationally efficient alternative.

Here, we combine the stochastic ab initio random structure search algorithm (AIRSS) with geometry optimisations based on interatomic potentials to systematically predict the structure of carbon clusters spanning a wide range of sizes. We first test the transferability and predictive capability of seven widely used carbon potentials, including classical (EDIP, ReaxFF, Tersoff, REBO-II, LCBOP-I, AIREBO) and machine-learning (GAP-20) potentials. Results are compared against an analogous cluster dataset generated via AIRSS combined with density functional theory optimizations. The best performing potential, GAP-20, is then employed to predict larger clusters in the nanometre scale, overcoming the computational limits of first-principles approaches. Our complete cluster dataset describes the evolution of topological properties with cluster size, capturing the complex variability of the carbon cluster family. As such, the dataset includes ordered and disordered structures, reproducing well-known clusters, like fullerenes, and predicting novel isomers."

Thermodynamically Informed Phase Space Exploration for Optimal Autonomous MLIP Dataset Building

Vincent Fletcher

University of Warwick, UK

Coauthors: 1: V.G. Fletcher, 2: A. P. Bartok, 3: L. B. Partay

1: Department of Physics, University of Warwick, Coventry, CV4 7AL

2: School of Engineering, University of Warwick, Coventry, CV4 7AL

3: Department of Chemistry, University of Warwick, Coventry, CV4 7AL

We present an optimal method of database generation for the training of machine learned interatomic potentials (MLIPs). Nested sampling is an unbiased Potential Energy Surface (PES) sampling technique that produces samples across all phases given no prior information. Since the accuracy of any MLIP depends on the underlying data it is trained on, and the data is required to undergo high cost ab-initio evaluation, selecting the fewest and most important data-points is a critical component in developing MLIPs efficiently. Samples generated by nested sampling form a sparse mesh of thermodynamically relevant points of the PES which creates a potent, low cost database that can be iteratively expanded through successive sampling runs. Based on the Atomic Cluster Expansion (ACE) we suggest a highly automated framework and, with our method, we reproduce fundamental properties of pure magnesium (vibrational and elastic properties, phase diagram, 0 K enthalpy curves) with remarkably small databases.

Session 3B (Thursday, 10:00-12:00)

Charge transport simulation using kinetic Monte Carlo with fast update rules

Roya Ebrahimi Viand

Fritz Haber Institute of the Max Planck Society, Germany

Coauthors: Roya Ebrahimi Viand, Sebastian Matera, Karsten Reuter,

Fritz Haber Institute of the Max Planck Society

The diffusion of particles in a solid material is often characterized by rare jumps between low-energy sites, which allows for the simulation of long-time dynamics using the kinetic Monte Carlo (kMC) methodology. In the case of charged particles, kMC can become costly because the long-range Coulomb interaction requires recalculating all process rates in every kMC step, despite the fact that each step introduces only local changes in the system's state. In this talk, we introduce update rules for, in general, anisotropic lattices, which require only two element-wise vector multiplications. This can be implemented very efficiently on modern compute architectures, and we discuss this using a newly developed kMC framework. We demonstrate the approach on a first-principles kMC model for Li^{+} transport in graphite in a high state of charge and investigate the dependence of mobility on the charge carrier density.

Exploring Triboelectric Charging Mechanisms with Density Functional Theory

James Middleton

University of Leeds, UK

Coauthors: James R. Middleton (1), Mojtaba Ghadiri (1), Andrew J. Scott (1)

(1) School of Chemical and Process Engineering, University of Leeds, Leeds LS2 9JT, UK.

The triboelectric effect, a phenomenon first documented by the ancient Greeks, remains a subject of scientific intrigue and debate. Despite extensive study, the fundamental mechanisms that drive this phenomenon—whether involving electrons, ions, or material nano-fragments—continue to be elusive, with no consensus among researchers. Recent technological developments in the field have facilitated the increased application of computational methodologies, particularly Density Functional Theory (DFT), which has opened new avenues for probing triboelectric charging mechanisms at the nanoscale. This study employs DFT to investigate the primary charge transfer mechanisms driving triboelectric interactions and to enhance the interpretation of experimental data. We apply DFT to gain insight into the fundamental molecular and atomic processes that underpin triboelectrification. Our approach includes evaluations of work functions across various crystal facets and polymers, which is considered an essential parameter for understanding how surface properties influence triboelectric charging. Additionally, we seek to explore the impact of environmental factors such as humidity and the presence of antistatic agents, both of which significantly affect triboelectric charging. We also investigate new materials suitable for triboelectric nanogenerators (TENGs), assessing their potential through analysis of their structural and electronic properties. Through this study, we probe the mechanisms of triboelectric charging and pave the way for innovative applications in energy harvesting and sensor technology.

Probing Structural Subtleties in Anti-Perovskite Solid Electrolytes

Karen Johnston

Durham University, UK

Coauthors: Karen E. Johnston,*1 George E. Rudman,1,2 Tavleen S. Attari,1 Theodosios Famprikis,3 James A. Dawson,2 Emma McCabe4 and Phoebe K. Allan5

1 Department of Chemistry, Durham University, Durham, UK

2 School of Natural and Environmental Sciences, Newcastle University, Newcastle, UK

3 Delft University of Technology, Delft, Netherlands

4 Department of Physics, Durham University, Durham, UK

5 School of Chemistry, University of Birmingham, Birmingham, UK

The rechargeable lithium-ion (Li-ion) battery is considered the technology of choice for energy storage in a wide array of portable electronic devices. However, its application is limited by its use of liquid electrolytes, which are known to pose a serious fire and safety risk. All-solid-state Li-ion batteries are attracting considerable attention as possible alternatives to conventional liquid electrolyte-based devices as they present a viable opportunity for increased energy density and safety. In recent years, a number of candidate materials have been explored as possible solid electrolytes, including garnets, Li-stuffed garnets, Li-rich anti-perovskites (LiRAPs) and thio-LISICONs. In particular, LiRAPs, including Li3−xOHxCl, have generated considerable interest, based on their reported ionic conductivities (~10−3 S cm−1).1,2 However, until recently, their lithium and proton transport capabilities as a function of composition were not fully understood.

Current research efforts have focused on the synthesis and structural characterisation of Li3−xOHxCl using a combination of high-resolution powder diffraction and variable-temperature multinuclear solid-state NMR spectroscopy with ab initio molecular dynamics. We will demonstrate that Li-ion transport is highly correlated with the proton and Li-ion vacancy concentrations. In particular, we will show that the Li ions are free to move throughout the structure, whilst the protons are restricted to solely rotation of the OH− groups. Based on these findings, and the strong correlation between long-range Li-ion transport and OH− rotation, we have proposed a new Li-ion hopping mechanism, which suggests that the Li-rich anti-perovskite system is an excellent candidate electrolyte for all-solid-state batteries.3 However, to fully understand the mechanism for conduction, multiple, complementary characterisation techniques are required. We will also present our latest findings regarding the room temperature structure of Li2OHCl, which has been the subject of considerable debate within recent years. We will demonstrate the complexities associated with elucidation of this structure and discuss the influence and impact of these structural findings on the ion conduction capabilities of this class of materials moving forward.

References

1. Y. Zhao and L. L. Daemen, J. Am. Chem. Soc., 2012, 134, 15042.

2. A. Emly, E. Kioupakis and A. Van der Ven, Chem. Mater., 2013, 25, 4663.

3. J. A. Dawson, T. S. Attari, H. Chen, S. P. Emge, K. E. Johnston and M. S. Islam, Energy Environ. Sci., 2018, 10, 2993."

Determining phase transitions in magnetic materials from first principles: Case study of MnAl

George Marchant

University of Warwick, UK

Coauthors: George A. Marchant (1), Laura H. Lewis, Julie B. Staunton (1)

(1) Department of Physics, The University of Warwick

The significant difference in energy product between the two primary classes of permanent magnets - ferrites and rare earth transition metals (RETM) - presents a challenge to fill the gap with an environmentally sustainable, RE-free magnet [1]. To this end, there is considerable interest in materials that are capable of forming L10-type layered structures, due to their lowered symmetry enhancing the material's intrinsic magnetic anisotropy (and in principle, coercivity and energy product). One such material is Mn_{55}Al_{45}, which has been shown to transform from the hexagonal, paramagnetic, solid solution ε-phase to the tetragonal, ferromagnetic, L10-type τ-phase under the influence of external magnetic fields. We investigate the ε-τ transformation using a hybrid diffusional-displacive (HDD) model - a computationally efficient representation of the possible pathways between the two states. Green's function-based density functional theory (KKR-DFT) is used to attain chemically disordered states, while elevated-temperature magnetism is treated at the level of electronic structure using disordered local moment theory. An atomistic model of the long range magnetic interactions is obtained from KKR-DFT data, while the model's thermodynamics are determined using magnetic nested sampling (MNS) to show the effect of a magnetic fields on phase stability. An introduction to MNS is given in terms of applications to toy models, to demonstrate its capability in determining the phase diagrams of magnetic interatomic potentials.

[1] Coey, J. M. D. ""Perspective and prospects for rare earth permanent magnets."" Engineering 6.2 (2020): 119-131.

[2] Ni, Chaoying, et al. ""L1 0 transformation characteristics of MnAl and FeNi under the influence of strain and magnetic fields."" Bulletin of the American Physical Society (2023).

Session 4A (Thursday, 14:30-15:40)

Machine learning of Gaussian basis sets for use in molecular applications

Grant Hill

University of Sheffield, UK

Coauthors: Shaun T. E. Donnelly, Robert A. Shaw, J. Grant Hill

Department of Chemistry, University of Sheffield, Sheffield S3 7HF

Gaussian basis sets have a large influence on both the accuracy and efficiency of molecular electronic structure calculations, but there have been few significant advances in their development over recent years. Recent work in our group has focused on using high-throughput calculations and machine learning to produce novel, molecule-optimised basis sets.

The open source Python package BasisOpt [1] has been used to optimise basis set exponents for molecules (rather than the usual optimisation for atomic energies), while the use of Legendre polynomials [2] reduces the dimensionality of the optimisation problem and avoids variational collapse. As a tool for automating the optimisation of basis sets, additional applications of BasisOpt will also be presented, including generation of auxiliary basis sets for use in density fitting, and reducing large basis sets into more efficient ones.

Analysis of data from the molecule-optimised basis sets leads to the use of machine learning (ML) to predict new basis sets for molecular applications. Results of ML basis sets will be presented and compared to existing alternatives, demonstrating that the approach is effective in producing accurate and efficient basis sets.

[1] Shaw, R. A.; Hill, J. G. J. Chem. Phys. 2023, 159, 044802.

[2] Petersson, G. A.; et al. J. Chem. Phys. 2003, 118, 1101–1109.

MESS: Modern Electronic Structure Simulations

Hatem Helal

Valence Labs, UK

Coauthors: Hatem Helal (1*), Andrew Fitzgibbon (2)

1. Valence Labs, London, UK

2. Graphcore, Cambridge, UK

*work initiated while employed at Graphcore

Electronic structure simulation (ESS) has been used for decades to provide quantitative scientific insights on an atomistic scale, enabling advances in chemistry, biology, and materials science, among other disciplines. Following standard practice in scientific computing, the software packages driving these studies have been implemented in compiled languages such as FORTRAN and C. However, the recent introduction of machine learning (ML) into these domains has meant that ML models must be coded in these languages, or that complex software bridges have to be built between ML models in Python and these large compiled software systems. This is in contrast with recent progress in modern ML frameworks which aim to optimise both ease of use and high performance by harnessing hardware acceleration of tensor programs defined in Python. We introduce MESS: a modern electronic structure simulation package implemented in JAX; porting the ESS code to the ML world. We outline the costs and benefits of following the software development practices used in ML for this important scientific workload. MESS shows significant speedups on widely available hardware accelerators and simultaneously opens a clear pathway towards combining ESS with ML.

Machine learning the electric field response of condensed phase systems using perturbed neural network potentials

Kit Joll

University College London, UK

Coauthors: Mr Kit Joll , Dr Philipp Schienbein , Dr Kevin Rosso, Professor Jochen Blumberger

1 Department of Physics and Astronomy and Thomas Young Centre, University College London, London, WC1E 6BT, United Kingdom

2 Pacific Northwest National Laboratory, Richland, Washington 99354, United States

3 Lehrstuhl für Theoretische Chemie II, Ruhr-Universität Bochum, 44780 Bochum, Germany

The interaction of condensed phase systems with external electric fields is of major importance in myriad processes in nature and technology ranging from the field-directed motion of cells (galvanotaxis), to geochemistry and the formation of ice phases on planets, to field-directed chemical catalysis and energy storage and conversion systems including supercapacitors, batteries and solar cells. Molecular simulation in the presence of electric fields would give important atomistic insight into these processes but applications of the most accurate methods such as ab-initio molecular dynamics are limited in scope by their computational expense. Here we introduce Perturbed Neural Network Potential Molecular Dynamics (PNNP MD) to push back the accessible time and length scales of such simulations. We demonstrate that important dielectric properties of liquid water including the field-induced relaxation dynamics, the dielectric constant and the field-dependent IR spectrum can be machine learned up to surprisingly high field strengths of about 0.2 V/Angstrom without loss in accuracy when compared to ab-initio molecular dynamics. This is remarkable because, in contrast to most previous approaches, the two neural networks on which PNNL MD is based are exclusively trained on zero-field molecular configurations demonstrating that the networks not only interpolate but also reliably extrapolate the field response. PNNP MD is based on rigorous theory yet it is simple, general, modular, and systematically improvable allowing us to obtain atomistic insight into the interaction of a wide range of condensed phase systems with external electric fields.

Session 4B (Thursday, 14:30-15:40)

Exploring the Pt(111)-Electrolyte Interface Under Applied Potentials with Ab Initio Molecular Dynamics

Clotilde Cucinotta

Imperial College London, UK

Chemistry, Imperial College London, UK

In this talk, I will discuss some complexities in the simulation of electrified interfaces at the nanoscale, focusing on the impact of applied potentials on their physicochemical properties. My approach is based on the development of highly realistic ab initio molecular dynamics models of charged electrode-electrolyte interfaces under bias. I will discuss recent advancements in modelling the double layer of the electrified Pt(111)-electrolyte interface, particularly in terms of its response to the applied electrode potential. This is achieved through the application of electrode-charging and potential control methodologies developed in my group [1-3]. If time permits, I will discuss how the how insights from molecular electronics can lead to a more sophisticated understanding of electrochemical phenomena.

References

[1] https://doi.org/10.1016/j.electacta.2021.138875

[2] https://pubs.acs.org/doi/10.1021/acs.jpclett.3c03615

[3] https://www.researchsquare.com/article/rs-3788305/v1

Ab-initio simulation of electron transfer processes of adsorbate-surface systems with ground-state density functional theory

Zsuzsanna Koczor-Benda

University of Warwick, UK

Coauthors: Zsuzsanna Koczor-Benda, Nils Hertl, Reinhard J. Maurer

University of Warwick

The interaction between electronic states of an adsorbate with the states of the underlying substrate is a decisive process for both natural phenomena, such as hydrogen formation on interstellar ice particles, and technological applications like plasmonic catalysis and ultrasensitive surface-enhanced Raman spectroscopy. The electronic coupling between molecular and metallic states enables electrons to be transferred across the interface, which can drive chemical reactions or enhance spectroscopic signals. Most often, in the so-called wideband limit, this coupling is assumed to be independent of energy. Going beyond this approximation requires calculation of state-to-state electronic couplings which is a challenging task. We investigate a projector operator diabatization (POD) method based on ground state density functional theory calculations that has previously been used for describing charge transfer processes between metal surfaces and weakly adsorbed atoms [1]. We utilise this method to calculate vibrational lifetimes of H atoms and CO molecules adsorbed on metallic surfaces such as Pt and Cu. Comparison to experimental data allows for an assessment of the applicability of this approach to molecular adsorbates and chemisorbed systems.

[1] Ghan, Simiam, et al. ""Interpreting ultrafast electron transfer on surfaces with a converged first-principles Newns–Anderson chemisorption function."" The Journal of Chemical Physics 158, 234103 (2023).

Computational modelling of oxygen-functionalized graphene: structure-property relationship and application in phosphate sensing

Natalia Martsinovich

University of Sheffield, UK

Coauthors: Negar Mansouriboroujeni, Xue Yong, Natalia Martsinovich

Department of Chemistry, University of Sheffield, Sheffield, UK

Graphene's excellent electronic and optical properties enable its use in a variety of applications. In particular, graphene's high electrical conductivity enables its use in sensors; however, selectivity is a challenge, which may be addressed by functionalizing graphene with various groups. We used theoretical modelling to investigate how oxygen functionalization affects the electronic properties of graphene and the interaction of graphene with phosphate, an essential plant nutrient. We collaborated with experimental partners to design graphene-based electrical sensors to detect phosphate [1].

Density-functional theory (DFT) calculations were used to model flat and curved graphene sheets containing epoxide and hydroxyl groups. We systematically studied a variety of arrangements of oxygen at different concentrations (1-12% O). Curvature was found to have little effect on the electronic properties of graphene: curved graphene remained semimetallic, with graphene-like optical absorption spectra. In contrast, oxygenation had significant effect on graphene's properties, with band gap opening in functionalised graphenes with oxygen content above 6%. This tuning of graphene's properties by chemical functionalization can open new applications in electronics and sensors. Next, we investigated adsorption of phosphate species on pristine and oxygenated graphene. Adsorption of phosphate affected graphene's band structure and caused distinct changes in pristine and oxygen-functionalized graphene sheets' electrical conductivities. In particular, our calculations predicted an increase in resistivity upon adsorption of phosphate on pristine graphene, which was experimentally verified by our collaborators [1]. These results suggest pristine and oxygen-functionalized graphene as promising materials for electrical sensors.

References

1. Yong, X. et al. ACS Applied Nano Mater. (2023) 10.1021/acsanm.3c04147.

Session 5A (Friday, 10:00-12:00)

Atomistic photodynamics simulations for solid-state energy materials

Federico Hernandez

University of Bristol, UK

Coauthors: Federico J. Hernandez and Rachel Crespo-Otero

School of Chemistry, University of Bristol. Bristol. UK

Department of Chemistry, University College London. London. UK.

Light-activated phenomena underpin applications in energy materials. This process takes place in the condensed phase, where the environment has an active role by either restricting the motions of the excited molecules or directly participating in the main excited mechanisms. A full understanding of these phenomena at the atomic level is required to optimize quantum efficiencies and aid the design of new materials with tailored properties. Excited state dynamics provides an insightful picture of light-activated processes at the intra- and intermolecular levels. However, computational studies of photophysics and photochemistry in the solid state is extremely challenging because of the substantial computational cost of high-fidelity calculations of chromophore excited-state dynamics in periodic systems. We developed a new approach combining a multiscale QM/QM' electrostatic embedding scheme with nonadiabatic molecular dynamics to model the photochemistry and photophysics of molecular crystals.

We test our approach studying the photochemistry of crystalline systems based on norbornadiene and azobenzene derivatives, top candidates for molecular solar thermal (MOST) applications, also known as solar thermal fuels. Solid-state MOST materials are particularly interesting as they can offer maximised energy densities and do not require encapsulation for applications as their liquid counterparts. However, the study of solid-state MOST systems is limited due to the added complexity arising from the solid environment. We modelled the excited sate dynamics from photon absorption predicting lifetimes and the corresponding quantum yields of the primary and degradation photoproducts formation with high level of accuracy.

Predicting and Interpreting Femtosecond X-ray spectroscopy

Thomas Penfold

Newcastle University, UK

Chemistry – School of Natural and Environmental Science, Newcastle University, Newcastle Upon-Tyne NE1 7RU, United Kingdom

Advances in experimental methodology aligned with technological developments, such as X-ray Free Electron Lasers (X-FELs) and High Harmonic Generation (HHG), has led to a paradigm shift in the capability of X-ray Spectroscopy to deliver simultaneously high temporal and spectral resolution. Importantly, the complex nature and high information content of this class of techniques means that detailed theoretical studies are often essential to provide a firm link between the spectroscopic observables and the underlying molecular structure and dynamics. For molecules in electronically excited states, these simulations must often go beyond the single nuclear configuration regime and the Born-Oppenheimer approximation.

Herein I will present some recent work on simulating and understanding ultrafast X-ray spectra using excited state simulations, including examples of predictions of ultrafast time-resolved experimental signals of organic and transition metal complexes. I will outline progress in electronic structure calculations and machine learning techniques to advance this field.

Mapping the landscape of electronic structure theory

Hugh Burton

University of Cambridge, UK

Coauthors: Hugh G. A. Burton, Antoine Marie, Sandra Saade

[1] Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, U.K.

[2] Laboratoire de Chimie et Physique Quantiques (UMR 5626), Université de Toulouse, CNRS, UPS, France

[3] Laboratory of Computational Science and Modelling, Institut des Matériaux, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland"

Multiple solutions are ubiquitous in electronic structure theory. They are found in linear and non-linear wave function approximations, provide indicators of strong electron correlation, and can be used to describe excited states. In this talk, I will outline recent progress in understanding wave function methods through the lens of energy landscapes. I will highlight the mathematical properties of the exact and approximate electronic energy landscape, and comment on their mutual connections. I will then outline how the break-down of wave function methods can be identified using symmetry breaking and how higher-energy solutions can be used to obtain an accurate and compact "state-specific" representations of electronic excited states. Molecular examples will highlight the advantages of state-specific approaches for high-energy, charge transfer, and double excitations, while also demonstrating the challenges and opportunities for future developments.

Surface-Specific Spectroscopy through Machine Learning

David Wilkins

Queen's University Belfast, UK

Centre for Quantum Materials and Technologies, School of Mathematics and Physics, Queen's University Belfast

In this talk, I show how machine-learning predictions of the energies and forces of atomistic systems, along with predictions of their responses to applied electric fields, can be combined with imaginary time path-integral methods to provide fully quantum-mechanical predictions of surface-sensitive sum-frequency generation (SFG) spectra. These calculations require that the polarizations and polarizabilities of the systems be accurately predicted; the former in particular requires some thought about how to build models for bulk polarizations. For this, I describe two approaches that render the polarization straightforward to learn.

Session 5B (Friday, 10:00-12:00)

Machine Learning Order Parameters in Atomistic Systems

Florian Dietrich

University College London, UK

Coauthors: Florian M. Dietrich (1), Michael A. Bellucci (2), Matteo Salvalaglio (1)

(1) Thomas Young Centre and Department of Chemical Engineering, University College London, London.

(2) XtalPi Inc, Cambridge MA

The efficient calculation of nucleation collective variables (CVs) is one of the main limitations to the application of enhanced sampling methods to the investigation of nucleation processes in realistic environments. In this work, we present the development of a graph-based model for the approximation of nucleation CVs that enables orders-of-magnitude gains in computational efficiency in the on-the-fly evaluation of nucleation CVs. By performing simulations on a nucleating colloidal system mimicking a multistep nucleation process from solution, we assess the model's efficiency in both postprocessing and on-the-fly biasing of nucleation trajectories with pulling, umbrella sampling, and Metadynamics simulations.

Moreover, we show how the model can be used to bias and obtain FESs in the space of powerful non-differentiable descriptors, which is not possible using traditional means, and how the computational efficiency gains can be even further expanded by learning multiple CVs at once. We believe this work lays the groundwork for simulating, notoriously hard-to-model nucleation processes of organic drug molecules.

1. Dietrich F. M., Advincula X.R., Gobbo G. Bellucci M. A., Salvalaglio M.. ”Machine Learning Nucleation Collective Variables with Graph Neural Networks”. J.Chem.TheoryComput. 2024, 20, 1600−1611"

Reversible simulation to train classical and machine learning potentials

Joe Greener

MRC Laboratory of Molecular Biology, UK

Molecular mechanics force fields are usually developed manually using quantum mechanical (QM) and condensed phase data, whereas machine learning potentials are typically trained on QM data alone. Ideally there would be automated methods that can reproducibly train potentials with any number of parameters to match a variety of data types. Here I present reversible simulation, a development of differentiable simulation, which calculates the gradients of observables with respect to force field parameters over a molecular simulation. Applications to parameterising molecular mechanics water models with various functional forms and a machine learning materials model will be shown. The approach is compared to existing ensemble reweighting methods and has the advantage of allowing fitting to dynamics data.

Efficient calculation of nucleation rates from equilibrium molecular dynamics and a state correspondence principle

Fabienne Bachtiger

University College London, UK

Coauthors: Fabienne Bachtiger [1], Aaron Finney[2], Linna Li [1], Matteo Salvataggio [1]

[1] UCL, [2] University of Liverpool

Simulating nucleation events of crystals from solution with the use of molecular dynamics (MD) is a challenging task. The common problems [1] are MD's inherent timescale limitations and the constraints imposed by system size, which give rise to the well-known finite-size effects[2]. Methods such as seeding MD [3,4] alleviate some of those issues, however, swarms of trajectories are often required to extract the desired rate parameters, rendering this methodology expensive. Furthermore, extracting the necessary rate parameters for low (and experimentally relevant) supersaturation is difficult as this usually requires unfeasibly large systems [5].

Here, we propose a new approach which builds on the seeding methodology, but where the stable equilibrium state is targeted instead of the unstable one[6]. We show that from a handful of simulations, a rate curve spanning a large range in supersaturation (including low and experimentally relevant conditions) can be computed. We embrace the finite size effects to find steady-state clusters and show their dual purpose in 1) computing the solubility and surface energy from a global fitting of the mean thermodynamic force and 2) by directly acting as the critical cluster in conditions where the background supersaturation reflects the one the steady-state cluster sits in. The latter relies on the state-correspondence principle[7].

Thus, when finding a steady-state cluster, we can simultaneously back-calculate the free energy barriers corresponding to the systems initial conditions (reflective of high supersaturation conditions) as well as directly calculate the barriers, as the steady-state cluster is also a critical one (reflective of low supersaturation conditions).

Importantly, using the state correspondence principle enables a direct calculation of attachment frequencies and, consequently, of absolute nucleation rates within the remits of Classical Nucleation Theory. We therefore present a cost-effective 2-for-1 approach in computing nucleation rate parameters using equilibrium MD simulations. We demonstrate the application of this approach to three distinct polymorphs of Olanzapine over a wide range of supersaturation. This case study would be impractical with more computationally expensive techniques, thus reinforcing its versatility and potential for broad use in the field.

[1] Finney, A. R.; Salvalaglio, M. Molecular simulation approaches to study crystal nucleation from solutions: Theoretical considerations and computational challenges. WIREs Computational Molecular Science 2023

[2] Salvalaglio, M.; Perego, C.; Giberti, F.; Mazzotti, M.; Parrinello, M. Molecular dynamics simulations of urea nucleation from aqueous solution.

Proceedings of the National Academy of Sciences of the United States of America 2015, 112, E6–E14

[3] P. Lamas, C.; R. Espinosa, J.; M. Conde, M.; Ramırez, J.; Montero De Hijes, P.; G. Noya, E.;Vega, C.; Sanz, E. Homogeneous nucleation of NaCl in supersaturated solutions. Physical Chemistry Chemical Physics 2021, 23, 26843–26852

[4] Zimmermann, N. E.; Vorselaars, B.; Quigley, D.; Peters, B. Nucleation of NaCl from Aqueous Solution: Critical Sizes, Ion-Attachment Kinetics, and Rates. Journal of the American Chemical Society 2015, 137, 13352–13361.

[5] Blow, K., Quigley, D., Sosso, GC., The seven deadly sins: When computing crystal nucleation rates, the devil is in the detail. Journal of Chemical Physics, 155 (4), 2021

[6] Li, L., Paloni, M., Finney, A., Barducci, A., Salvalaglio, M., Nucleation of biomolecular condensates from finite-sized simulations. Journal of Physical Chemistry Letters, 14(7), 1748-1755, 2023

[7] Grossier, R.; Veesler, S. Reaching one single and stable critical cluster through finite-sized systems. Crystal Growth and Design 2009, 9, 1917–1922"

Coatings for Corrosion Protection: Insights from Molecular Simulations

Flor Siperstein

University of Manchester, UK

Coauthors: Kieran Harris, Charlie R. Wand, Peter Visser and Flor R. Siperstein

1. University of Manchester, UK
2. University of Exeter, UK
3. AkzoNobel, Sassenheim, The Netherlands
4. University of Manchester, UK

Protecting metal assets from corrosion is important from a safety, economic and sustainability perspectives. Coatings are a common approach to protect metals from corrosion, and understanding the interfacial properties where the organic coating is in contact with the solid substrate, can aid in the development of coatings with better performance.

Modelling and simulations can provide insight into problems that experiments cannot explore, complementing experimental results. In epoxy-based coatings, the significance and impact of surface-induced segregation of the coating during the cure has proved difficult to investigate through experimental means alone.

We have created a robust workflow to model the crosslinking reactions of various precursors in the presence of a surface, overcoming challenges of previous methods related to the implementation of the atomistic reaction and the definition of reacting units.

As a model epoxy-amine system, diglycidyl ether of bisphenol A (DGEBA) and m-xylylenediamine (MXDA) have been used. We find an enrichment of MXDA at the solid interface, resulting in a reduced crosslink density at the interface compared to the bulk. Further, the depletion of MXDA from the film beyond the contact layer leads to the low crosslink density region extending further into the film than the amine enriched contact layer. This results in distinct regions of excess amine and excess epoxy close to the surface.

The consequences in performance of this phenomena require further investigation, but it is possible to envision advantages and disadvantages. On one hand, the low crosslinking at the interface leaves uncoordinated amine groups that can form bonds with the metal oxide surface, providing adhesion. On the other hand, the reduction of crosslinking near the surface and in the film may result in the pathways allowing water, small molecules, and ions to travel through the coating, reducing the corrosion protection performance.